

# Turning the Accountability Tables: Ten Progressive Lessons from One 'Backward' State

For too long teachers have been forced into a defensive posture, protecting their professionalism and their students' learning from the accountability hawks who know little about teaching and learning, Mr. Gallagher asserts. It's time to turn the tables.

BY CHRIS W. GALLAGHER

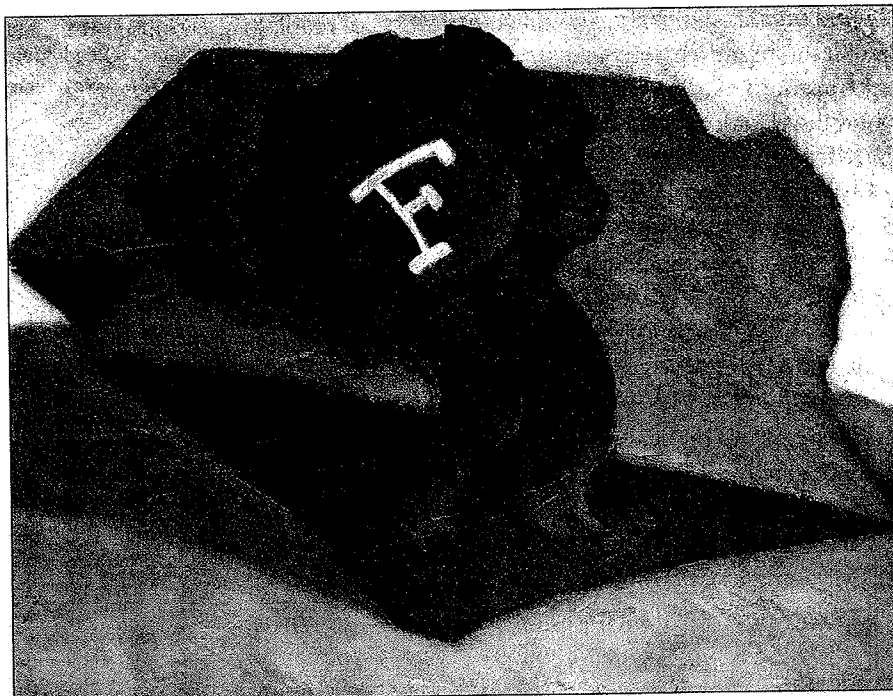
**E**VERY YEAR, like an anxious teenager waiting for that all-important SAT or ACT score to arrive in the mail, I eagerly await the release of The Report. I spend days, weeks, months, checking the Web and praying to the accountability gods. I send up my plaintive cry: *please, let us do just as well as we did last year.*

And, I am happy to say, I am never disappointed.

Every year, Nebraska gets an F.<sup>1</sup>

*Meanwhile . . .* the kids in Palmer do Community Math, skillfully solving complex real-world problems they have solicited from their neighbors and parents. The kids in Cedar Bluffs do the "Platte Attack," writ-

CHRIS W. GALLAGHER is an associate professor of English, University of Nebraska, Lincoln. He presented an earlier version of this article as the keynote address for the fall meeting of the Western Pennsylvania Writing Project, October 2002.



ing wonderful sandbar poetry and keeping science journals on the banks of the river that runs through their rural town. The kids here in Lincoln develop math portfolios, reflecting on and documenting their

learning every day. The kids in Heartland build immigrant trunks and capably present their family heritage projects to a full auditorium every year.

An F means that Nebraska con-

tinues to buck the high-stakes, test-'em-'til-they-drop mentality. It means that in Nebraska, assessment continues to be driven by instruction, rather than the other way around. It means that in Nebraska, as one teacher aptly puts it, having standards "does not make us all 'standard.'"<sup>2</sup>

## ACCOUNTING AND ACCOUNTABILITY

The F comes from *Education Week's* annual *Quality Counts* (QC) report. This report never fails to put me in mind of a certain Lily Tomlin quip. "I worry," she said, "that whoever thought up the term 'quality control' thought if we didn't control it, it would get out of hand."<sup>3</sup> And a good way to control things, as we all know, is to count them. One thing we can say about *Quality Counts*: it's aptly named.

For the past six years, *Education Week* has published this 50-state report card on education. The QC system (yes, right, the QC system) has several components, but here I want to focus on "standards and accountability," the category in which Nebraska, my home state, consistently and gloriously earns an F.

Each state's grade for "standards and accountability" is based on several elements. Fifteen percent of the grade is based on whether or not the state has adopted standards. Twenty-five percent is based on whether or not those standards are clear and specific (a yes/no question, apparently). Two percent is based on whether or not the state participates in the National Assessment of Educational Progress.

Now here is where the QC system gets interesting (if a little baroque). Twenty-eight percent of the grade is based on assessment. Within this category, states get points for having each of five types of assessment in each subject at each grade level. States are also awarded points for having aligned criterion-referenced tests in each subject at each grade level. Elegant formula: more tests = more points. The remaining 30% of the grade is based on accountability, with states receiving points for having report cards, ratings, rewards for high-performing schools, and assistance and sanctions for low-performing schools. Another elegant formula: higher stakes = more points.

Now, it is a good idea to hold state accountability systems accountable. In fact, I have spent the past three years leading a team of researchers who are studying one state accountability system in particular. But I have found few responsible examples to draw from in this area.<sup>4</sup> In fact, as a recent RAND report notes, we have

little research to suggest that current accountability systems do more than artificially raise tests scores.<sup>5</sup> We do not even know that the kinds of sanctions promoted by QC have salutary effects on low-performing schools. In fact, we should wonder whether these accountability systems generally, which are based on manufacturing models of productivity, still fit with our educational goals. The practices informing these models — standardization, ranking, top-down control — historically have been tied to the goal of sorting students for the work force, not teaching all children to a high level.

According to RAND, what we are seeing in states leading the accountability charge — Texas and Florida, for instance — should also give us pause. To wit:

- students are becoming discouraged and competitive, rather than cooperative;
- teachers are becoming alienated and are narrowing curricula to teach to tests;
- administrators are becoming number-obsessed and are overspending on tests and other purchases peripheral to teaching and learning;
- policy makers are becoming misinformed about learning and focused on the politics of perception; and
- communities are also becoming misinformed, as well as distrustful of schools and teachers.

The RAND report, to be fair, does pose several *potential* benefits of high-stakes testing. For instance, such testing *could* motivate students and teachers to work harder, offer administrators helpful information about their programs, and aid policy makers in analyzing the effectiveness of policies. Any reader of the report, however, will recognize the profound, if implicit, skepticism beneath its surface evenhandedness. In any event, as such research reports tend to do, this one ends with a sober call for . . . more research. Which is fine, but in the meantime, students, teachers, and communities are being subjected to a powerful accountability movement that gets education all wrong.

## REFORMING REFORM

The fundamental problem is that these reformers focus — to borrow a useful formulation from Linda Darling-Hammond's wonderful book *The Right to Learn* — on *designing controls* rather than *developing capacity*. In other words, instead of promoting and investing in the expertise of teachers and trusting them to do their job, most state systems focus their resources on building remote-control systems, in which "experts"

— administrators, policy makers, politicians, curriculum designers, textbook companies, or testing firms — set and measure the educational agenda from afar.<sup>6</sup>

The fatal flaw in this approach is not hard to see, but, as such historians as Lawrence Cremin, Larry Cuban, David Tyack, and Darling-Hammond tell us, it has haunted the history of U.S. education reform.<sup>7</sup> The mistake is treating school reform as a technical problem, not a people problem.<sup>8</sup> Reformers seem to “forget” again and again that institutions are made up of people, and these people constitute a local culture that must be engaged if long-term change is to be sustained. Today’s top-down reformers prove themselves no different when they locate the engine of reform outside schools. That is why they will not, in the long run, be successful.

That is the good news. The bad news is that the long run is a *long run*. If the history of education is any guide (and if I may unceremoniously switch metaphors), this storm *will* blow over. But — fueled by the force of federal policy, state acquiescence, and, most dangerous of all, unthinking common sense — it will not do so any time soon.

And there is more mixed news: we will continue to see grassroots efforts such as the parent/teacher/student anti-high-stakes-testing alliances that have sprung up across the country, as well as experimental progressive schools that almost always show extraordinary learning gains.<sup>9</sup> But these important local efforts will continue to be isolated and will take place primarily in affluent communities where resources and networks are sufficient to argue successfully for waiver status.<sup>10</sup>

This “special case” mentality points to a critical limitation of such grassroots efforts: they tend not to reckon with institutional structures, preferring to argue for a space outside them for targeted (typically already privileged) populations. They miss a crucial feature of progressive education, as John Dewey imagined it almost a century ago. “It is the aim of progressive education,” he wrote, “to take part in correcting unfair privilege and unfair deprivation, not to perpetuate them.”<sup>11</sup> That is why Dewey urged progressives to operate on the principle that we must provide *all* students with the quality of education the “best and wisest” parent wants for her child.<sup>12</sup>

Of course, this is a rather tall order, and while it has made its way into the No Child Left Behind rhetoric, I am not sure any of us has a handle on what it would actually take to mount such an effort. This may be why so many of us tend toward despair when our think-

ing turns to the big picture. We tend to view institutions and systems as nameless and faceless monoliths operating outside the bounds of our comprehension, never mind our efficacious action. Even — or maybe especially — those of us within the system feel powerless to influence its machinations.

But the truth, as those historians remind us, is that institutions and systems *are* changeable, and in fact are constantly changing. As individuals, we may not be able to make grand, sweeping changes in the bureaucracies in which we work. But we should recognize that these bureaucracies are reconstituted daily through *human* activity.

The present educational and political climate, I realize, does not support this kind of thinking; in fact, it is designed to suppress it. As states rush to comply with new federal mandates, few observers are moved to wonder whether those mandates are good for kids and teachers in the first place. The voices of those who are so moved are drowned out in a veritable chorus of resignation: *That’s just the way it is — the feds have spoken, and it’s hopelessly idealistic to imagine alternatives.*

But as all good teachers know, *imagining alternatives* is at the heart of teaching and learning. Conversely, they know that nothing so deadens the teaching and learning process as the abject belief that no alternatives exist. Absent a sense of possibility, why explore, wonder, analyze, critique, experiment, inquire — indeed, why learn at all?

I know that this very commitment — to help students open up to worlds of possibility — is what keeps so many fine teachers in the profession. It is why they suffer the slings and arrows of outrageous demands, develop strategies of subtle subterfuge, or participate in the noble but frustrating and unglamorous task of tinkering with tests. And I confess, when I begin thinking in this direction, I begin to get all warm and fuzzy. But I don’t want to romanticize teachers here. Romanticizing the work of teachers is what allows newspapers to run “My Teacher, My Hero” features while supporting plans like high-stakes teacher testing or the revocation of tenure. And it is what allows state departments of education to hand out Golden Apple awards with one hand and prepackaged curricula with the other. In short, it is all too often an avoidance tactic. If we praise teachers enough, we not only do not need to pay them much, but we also do not need to trust their professional judgment, invest in their ongoing professional development, or create conditions of possibility for full and rewarding professionalism.

And here is the point our present accountability hawks *and* our grassroots activists miss: *systematically* creating conditions of possibility for knowledgeable professionals to practice their art adaptively must be a crucial feature of any lasting education reform. Good teaching here and there is not enough. Securing meaningful learning opportunities for our own students or decent working conditions for ourselves and our own colleagues is not enough. We — educators, parents, and other stakeholders who want meaningful education for all students — need to work together for structural change as well.

A number of networks and organizations devoted to structural change have emerged in recent years, including the Coalition of Essential Schools; the National Coalition of Education Activists; the School Development Program; the National Network for Educational Renewal; Accelerated Schools; the National Center for Restructuring Education, Schools, and Teaching; and FairTest.<sup>13</sup> These networks and organizations provide useful resources and heartening examples of teachers and communities working together for change. They help us analyze the seemingly inscrutable systems in which we work. They remind us that the current accountability movement is neither natural nor inevitable. In short, they help us remember that alternatives *are* possible.

#### AN ALTERNATIVE: NEBRASKA'S STARS

At the same time, we are seeing very little innovative work *at the state level*, where the dominant posture, as I have suggested, is resigned compliance. Those who are doing the good work of these organizations and alliances find themselves swimming against the tide of top-down state requirements and demands designed to do the bidding of an even more top-down federal government.

One state, at least, is resisting the rush to comply, attempting to ensure that whatever changes it makes in the way it educates its children are pedagogically sound and professionally responsible. Under the courageous leadership of Commissioner Douglas Christensen, Nebraska is carefully balancing federal demands with state needs and goals. This is an ongoing process, of course, but it is clear that this state is not willing to acquiesce unthinkingly to the anti-teacher, public-school-bashing testing frenzy that has gripped the nation, from our President on down.

Call Nebraska backward. Call it slow. Or, as some

observers do, call it an emerging leader in assessment and education reform.<sup>14</sup> Its unique, local-control approach to school renewal may emerge from a deeply conservative tradition, but make no mistake: compared to what we are seeing in other states, Nebraska is a veritable hotbed of progressive educational activity.

In fact, as coordinator and principal investigator of the comprehensive evaluation of Nebraska's accountability system, and as someone who has spent much of the past three years listening to teachers describe their work, I've come to think that the experiments undertaken in this state have important implications for those of you who teach elsewhere. I understand, of course, that Nebraska's system is no silver bullet and that it cannot be replicated everywhere. On the contrary, perhaps the chief virtue of this system is that it was *not* prepackaged, but instead grew out of the unique circumstances of its context. I offer it here, then, not as a *model* to be imitated, but rather as an *example* — an alternative. I cannot describe Nebraska's standards, assessment, and accountability system in great detail, but broad strokes should suffice to give you a sense of it.

Hoping to preserve local decision making and avoid the pitfalls they have witnessed in states that have adopted high-stakes standardized testing, state policy makers have designed a *system of local assessments*.<sup>15</sup> In other words, rather than develop a single state test, as most states are doing, Nebraska has built a system that uses a variety of assessments, including classroom-based assessments, for reporting. The genius of this system is that it uses formative assessment for summative purposes.

In fact, STARS (School-based Teacher-led Assessment Reporting System) is unique in several ways. Specifically, it

- is a system of local assessments, not a state test;
- promotes a balanced approach to assessment, using multiple measures;
- aims for equivalence without sameness;
- involves evaluation of achievement and of assessment quality;
- uses classroom-based assessments for reporting; and
- includes no high-stakes testing.

The process looks like this. Districts must first adopt state or local standards in language arts, math, social studies, and science. The state has set a schedule requiring districts to report on one of these subject areas per year at grades 4, 8, and 11. In September, districts submit an assessment plan, and the Nebraska Department

of Education (NDE) provides feedback. Some districts elect to build their assessment programs around existing standardized tests — though researchers have found that only a handful of the state standards are covered by such tests. In these districts, criterion-referenced district tests and often classroom-based assessments are used to cover the remaining standards. Other districts use only local tests and, in a handful of cases, only classroom-based assessments. In any event, districts develop, administer, and score their local assessments throughout the school year. Then, in June, they submit a District Assessment Portfolio, which includes information on student performance on the standards, as well as on the assessments used to measure that performance, including sample assessments. These portfolios are reviewed and rated for *both student performance and assessment quality* by two groups of independent experts using six criteria:

1. Assessments reflect state or local standards.
2. Students have an opportunity to learn the content.
3. Assessments are free from bias or offensive language or situations.
4. The level is appropriate for students.
5. There is consistency in scoring.
6. The mastery levels are appropriate.

The dual ratings become the basis for the *State of the Schools Report*, an exhaustive archive of information on school conditions, demographics, performance, and assessment quality, available in English and Spanish on NDE's website. This information is abstracted in a state report card and widely disseminated in newspapers across the state.

Results from the first year of reporting (2000-01) were encouraging. Assessment quality ratings showed that the majority of assessments — 65% — were either exemplary or very good, and only 8% were rated unacceptable. Student achievement results were roughly the same, with 60% in the exemplary or very good categories and 17% unacceptable. In nearly everyone's view, this is a good start but leaves much to be done.<sup>16</sup>

## CHARTING STARS

In order to monitor the effects of STARS, the Nebraska Department of Education contracted with the University of Nebraska for an independent, comprehensive evaluation. I was asked to lead the project because of my background, which included previous research on assessment, work with Nebraska teachers,

and experience in the testing industry.

I began by assembling a cadre of researchers at the university. We decided to employ a variety of qualitative and quantitative research methods, including surveys, interviews, observational research, and discourse analysis. For the first year of the project, we focused on

- teachers'/administrators' perceptions of STARS;
- professional development and capacity building;
- effects of STARS on classroom practice, curriculum, and school climate; and
- leadership challenges posed by the system.

So what did we find? Our data have proved both heartening and sobering for the state. There is certainly a great deal to celebrate about STARS, but there are also serious challenges and a great deal of work to do yet.<sup>17</sup> Among strengths, we found that STARS has considerable, if tentative, support among teachers and administrators; that it is sponsoring high-quality professional development for teachers; that it is promoting collaborative, critical examination of curriculum; and that it is leading teachers to more informed classroom practice. These seem to us key ingredients in developing a sustainable approach to school renewal.

At the same time, our research revealed some serious challenges to sustainability. Specifically, we found some resentment of the whole process, especially on the part of long-time teachers and administrators, who resented the time demands it imposes. We also found insufficient teacher participation in STARS across grades and curricula; lack of confidence in teacher expertise; and inadequate engagement of stakeholders outside the schools, especially parents. Each of these challenges, if left unchecked, would spell doom for a school-based, teacher-led system.

These findings are proving useful to Nebraska policy makers, who are currently drafting an action plan in response to our report. But I am convinced the STARS system and our research also offer important lessons for teachers elsewhere who want to do more than wait out this long storm, who are willing to embrace their professional responsibilities, and who want to put pressure on their states to do the right things rather than do things right.<sup>18</sup> I offer the following in the energizing spirit of seeking alternatives and against the enervating spirit of securing acquiescence.

## TEN PROGRESSIVE LESSONS FROM ONE 'BACKWARD' STATE

1. Teachers should be regarded as leaders of reform, not

*impediments to it.* The long road of educational history is littered with failed reform agendas that were designed *for or around* teachers, never *with* them. In Nebraska's teacher-led system, by contrast, local educators are the locus of decision making: they design and administer curriculum and instruction, and they *develop, administer, and score* assessments that measure student learning. These assessments serve teachers' classroom needs as well as fulfill reporting requirements. Perhaps this accounts for the many teachers who told us that they were actually energized by their participation in STARS; as one educator put it, the system "validates [their] professionalism."

2. *Accountability systems must focus on developing capacity, not controls.* In Nebraska, as everywhere, too many resources are allocated to bureaucratic activities peripheral to teaching and learning. However, the state is moving toward a system in which the primary and direct targets of funding and expertise are local educators. Instead of focusing energies and money on a remote-control apparatus, the state invests heavily in ongoing teacher learning. As a result, we found a high level of professional development among teachers, especially in the area of assessment literacy. Here is one teacher's description of teacher learning in her school:

We wouldn't be teachers if we didn't know how to teach. But actually breaking things apart and really using effective strategies, I think, is the key when you're teaching. And that might not have been a real strength for some teachers, but now we're making it a part of our whole school, that process where teachers really are learning the best practices and implementing them in their classrooms.

This school, clearly, has become a *learning organization*, in which all participants, including teachers, become learners.<sup>19</sup> This is a goal for all Nebraska schools.

3. *Accountability systems must foster commitment, not compliance.*<sup>20</sup> This idea follows from the previous principle. If the state wishes to sponsor the learning of teachers as well as students, then it must promote *will*ing, *internal* accountability first. In Nebraska, teachers told us that the resources for professional development certainly help but what really motivates their active participation is a combination of trust in their professional judgment and the belief that this work will be beneficial for students. These teachers operate from a primary commitment to their professional work,

not to pleasing authorities. I am struck by the simple elegance of what one teacher told us: "What we're doing makes sense to us. And it seems to satisfy them, so . . ."

4. *Accountability systems must promote integration of school improvement and accountability efforts.* This is the third principle writ large: when schools and districts operate from a compliance perspective, their work on accountability tends to be cynical and mechanical. Conversely, when they operate from a commitment perspective, they see accountability as locally meaningful — as a route, in short, to school improvement. Here is how one superintendent describes his district's work on STARS:

We're saying to the state, "We're not doing this for you. We're doing this for us." And ultimately, that's the way it's supposed to be. So we feel really good about where we're going and what's happening. It's a lot of work, and the teachers understand that we don't just do this one time. We're going to continue every year looking at our results, changing our curriculum, etc.

It is significant that this superintendent and the teachers in his district reported (separately) that they would be doing this work whether or not the state required it.

5. *Accountability systems must risk complexity rather than demand simplicity.* Unfortunately, current reformers — those trumpeters of creative, critical, and higher-order thinking — have set a bad example in their own work. Nothing is simpler — or less creative — than the idea that all students should be given the same test and that the results of that test should be used to rank students, teachers, and schools. If schooling were a competitive sport, as Alfie Kohn reminds us, then such tactics might make sense — after all, the goal would be to declare winners and losers.<sup>21</sup> But what about this "no child left behind" idea? Students *will* be left behind in the high-stakes, standardized testing system; that's the whole *point*.

But make no mistake: Nebraska's alternative approach *is* complex. There is nothing simple about compiling or rating a District Assessment Portfolio. There is nothing easy about designing assessments that are good for students *and* meet rigorous quality criteria. But then, there is nothing easy about teaching and learning. They are more art than science, more messy human process than neat mechanical skill. The most

devastating effect of curriculum-narrowing standardized tests, in my view, is the sacrificing of spontaneity, creativity, experimentation, and risk-taking — all the “higher-order” capacities that we say we want for our students but that cannot be tested by multiple-guess or even short-essay testing. Nebraska teachers agree — and they take every opportunity to remind us — that protecting and enhancing their students’ ability to learn in their own diverse, complex ways remains among their top priorities.

6. *Accountability systems must include all students.* We have heard reports of the reprehensible practice of educational triage, in which only “marginal” students — those who stand to make significant gains on test scores from year to year — are given significant attention.<sup>22</sup> And we know schools play a shell game in which the trick is to hide special education students and English-language learners, either by exempting them from assessment and reporting or by engaging in another disturbing practice: educational gerrymandering (herding special populations into one school in order to raise the scores of the others). We also know that the best teachers are far more likely to be assigned to already high-achieving schools than to those where the students have more pronounced needs. Unfortunately, as study after study shows, qualified, highly skilled teachers are unevenly distributed along racial and class lines both within and across schools, districts, and states. Couple this with the well-documented inequity of school funding formulas, most of which continue to be built on local tax bases, and we have more than what President Bush is fond of calling “the soft bigotry of low expectations”; we have the hard bigotry of systemic inequity.

In Nebraska, the state department of education is leaning hard on *contextualizing* results, changing the question from “How high are this school’s scores?” to “How well is this school addressing its challenges and building on its strengths?” Our commissioner continually reminds the media and others that, as he puts it, “rank-ordering schools undermines everything we want to do.” The purpose of accountability is to generate community conversations aimed at improvement from *wherever the school starts*. And this does *not* mean generating arbitrary or unrealistic quantitative measures of “adequate yearly progress”; it means charging schools with the responsibility of setting ambitious but reasonable goals, developing plans for achieving the goals, and publicly measuring their progress toward them. It also means going beyond simplistic ideas about “op-

portunity to learn.” We need to focus not only on *whether* content was presented, but also on *how* it was presented. Teaching must be responsive to what students already know and how they think if we are to leave no child behind.

7. *Accountability systems must also leave no teacher behind.* The best way I know of to leave no student behind is to leave no teacher behind. If all teachers share responsibility for helping students reach high standards, if all teachers have ample and ongoing professional development, and if all teachers have a hand in shaping school improvement goals and designing practices intended to achieve those goals, then far fewer students will fall through the cracks. As I have said, Nebraska has work to do in this area; too much of the burden for assessment and reporting is falling on the shoulders of reporting-grade teachers. However, our researchers have also seen the reverse: districts in which there is broad “buy-in” among teachers. Here, we saw extraordinary levels of collegiality, engagement, and self-efficacy among teachers. Consider this teacher’s description of her school: “You could walk in, and the K-12 staff is knowledgeable about the whole assessment process. There’s nobody left out in the woods, from the business teacher to the tech teacher, and grade-level teachers have really talked. They know what everyone else is doing.”

Nebraska has committed most of its professional development resources to the creation of learning teams, rather than to inservices or more traditional training, and the results in many districts have been remarkable: broad teacher buy-in, a focus on team-building, and curricular integration. While the state is a long way from involving every teacher in school improvement, a recent study finds that more and more Nebraska schools are becoming collaborative learning environments.<sup>23</sup>

8. *Accountability systems must engage all stakeholders.* Let us take this one step further. While teachers and students are the most crucial actors in school reform, education is finally best viewed as a whole-community responsibility. “Informed conversations and informed decisions,” our commissioner has said, “are the heart and soul of democracy.” Fostering informed conversations begins with clear communication. But such expectations not only should be *communicated* to all stakeholders — within and beyond the system — but also should be *shaped* by them. I do not mean that those outside the schools should mandate and micro-manage what happens within them; rather, just as an effective architect responds to the local culture and ecol-



ogy, so too does the effective teacher. The professional standard here would not be for teachers to acquiesce to every idea that is suggested to them, but rather to engage in dialogue with a wide array of stakeholders as they design their curricula and teaching. In return, community members would expect, and trust, teachers to make decisions informed by current professional standards and a concern for the welfare of the local community, beginning with students. As one teacher we interviewed told us, "We are saying, 'This is our philosophy. This is what we believe assessment is about.' And we're getting that information to parents and to our schools, administrators, teachers, and students themselves."

9. *Accountability systems must keep pedagogy — teaching and learning — at the center.* This is of course what it's all about — the "heart of the matter," as Beverly Falk calls it in her book of that name.<sup>24</sup> We are fortunate in Nebraska to be working with measurement experts who do not see technical quality and classroom assessment as mutually exclusive and who understand that, if we must sometimes choose between an assessment of pristine technical quality and what's best for a certain group of students under particular circumstances, then we must choose the latter. They know that test items can be pretested and determined to be perfectly valid and reliable without having any educational value whatsoever. They have worked hard with the state department to develop model procedures for ensuring the technical integrity of various kinds of assessments, including classroom-based assessments. This is crucial because it allows the state to maintain its focus on teaching and learning, rather than searching for "the perfect test."

10. *Accountability systems must promote high-impact, not high-stakes, assessment.* High-stakes accountability is really a form of accounting, in which the winners — those with the most points (highest scores) — are rewarded, and the losers — those with the fewest points (lowest scores) — are punished. Sadly, it is also based on a flawed learning theory. Any teacher who relies solely on extrinsic motivation — the wielding of carrots and sticks — is committing professional malpractice. But that's exactly what high-stakes accountability systems do.

In Nebraska, policy makers, administrators, and teachers have at least begun to shift their mindset from "How will this assessment help us keep score?" to "What consequences will this assessment have on teaching and learning?" They have taken seriously recent studies —

many of which are summarized in Paul Black and Dylan Wiliam's article in the October 1998 *Kappan* — that formative, classroom-based assessment promotes student learning.<sup>25</sup> To borrow language from Richard Stiggins, the emphasis is not only on assessment *of* learning, but also on assessment *for* learning.<sup>26</sup> As one educator emphasized to us, "STARS really does individualize instruction for students." In other words, because her assessments are locally designed and grow from the projects she has her students develop, this teacher is able to be responsive and innovative in her classroom. That's "high-impact" assessment.

## TURNING THE TABLES

As I wind up this article, let me presume for the moment that you are drawn to the ideas I have been sharing with you. Even if you do like what I have said, you might be thinking, "Okay, but tell *them*: the policy makers and bureaucrats who run the system." It may seem, in other words, that in targeting this article — which deals primarily with policies and structural changes — to teachers, I am focusing on the wrong crowd.

But this is precisely my point: we already have quite enough of people like *me* (university types) talking to people like *them*. It is time for those who really do and should control the system — teachers — to be telling the rest of us what's educationally best for their students. And it is time for teachers to turn the accountability tables on the reformers, keeping *them* accountable by asking the kinds of questions that really matter. It is in this regard that I hope these lessons might be helpful. Instead of asking, "Does this system raise test scores?" and "Is it efficient?" we'd be better off asking questions that correspond to Nebraska's 10 principles of sound accountability.

- Does this system regard teachers as leaders?
- Does this system focus on capacity rather than controls?
- Does this system foster commitment and not mere compliance?
- Does this system promote integration of accountability and school improvement?
- Does this system risk complexity rather than demand simplicity?
- Does this system really include *all* students?
- Does this system engage all teachers?
- Does this system engage all other relevant stakeholders?
- Does this system keep pedagogy at its center?



• Does this system encourage high-impact, not high-stakes, assessment?

Teachers are just the people to be asking such questions. Teachers have historically proved to be the chief impediment to lasting school reform. (This, by the way, should be a point of pride; I tell policy makers that the best evidence we have that teachers are trustworthy is that they have resolutely refused to jump on any old bandwagon that happens by.) That is why teachers are the key to genuine change, genuine renewal. No matter how draconian reformers' controls might be, they won't get anywhere, ultimately, without the support of those who stand at the point of contact with students and communities.

But more than that: for too long, teachers have been forced into a defensive posture, having to protect their professionalism and their students' learning from accountability hawks who know little about teaching and learning. But who is keeping the accountability hawks accountable? Who is ensuring that their goals and practices are good for students and teachers? Who is insisting that their proposed sanctions, for instance, improve learning? To be sure, we — the public, higher education, publications such as the *Kappan* — should all be doing this work together. But as Nebraska educators have taught me, teachers — the professionals who stand at the point of contact with students and communities — are the best positioned to help us ask better questions and to turn the accountability tables.

1. Editorial Projects in Education, *Quality Counts 2003*, available at <http://www.edweek.org/sreports/qc03>.

2. Ellen Kohtz, "Reflection," in *Authentic Evaluation: Nebraska Teachers Design Assessment Through Active Classroom Learning* (Lincoln, Neb.: Goals 2000 Assessment Project, 2000), p. 8.

3. Lily Tomlin, quoted in Sandra Murphy and Barbara Grant, "Portfolio Approaches to Assessment: Breakthrough or More of the Same?," in Edward M. White et al., eds., *Assessment of Writing* (New York: Modern Language Association, 1999), p. 284.

4. An exception is a policy brief released by the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). The brief usefully calls for evaluation of accountability systems, and its proposed standards — on system components, testing standards, stakes, public reporting formats, and evaluation — offer some valuable guidelines, such as using multiple measures of student learning and designing appeal procedures to contest rewards and sanctions. At the same time, the brief has limited utility in asking more fundamental questions, such as "Why an accountability system in the first place?" and "Who has a voice in its development and enactment?" See Eva L. Baker et al., "Standards for Educational Accountability Systems," CRESST, Policy Brief No. 5, Winter 2002, available at [http://cresst96.cse.ucla.edu/products/policybriefs\\_set.htm](http://cresst96.cse.ucla.edu/products/policybriefs_set.htm).

5. Brian M. Stecher and Laura S. Hamilton, "Putting Theory to the Test," *RAND Review*, Spring 2002, pp. 17-23.

6. Linda Darling-Hammond, *The Right to Learn* (San Francisco: Jossey-

Bass, 1997). Like Deborah Meier, I use "trust" not to mean blind faith, but instead an *earned* trust in which teachers and schools are expected to demonstrate wise professional judgment. See Deborah Meier, *In Schools We Trust* (Boston: Beacon Press, 2002).

7. Lawrence Cremin, *The Transformation of the School* (New York: Vintage, 1961); Larry Cuban and David Tyack, *Tinkering Toward Utopia* (Cambridge, Mass.: Harvard University Press, 1995); and Darling-Hammond, op. cit.

8. See Peter H. Johnston, *Constructive Evaluation of Literate Activity* (New York: Longman, 1992).

9. See [www.fairtest.org](http://www.fairtest.org) for information on grassroots anti-testing alliances.

10. See Darling-Hammond, op. cit.

11. John Dewey, *Democracy and Education* (1916; reprint, New York: Free Press, 1966), p. 120.

12. John Dewey, *The School and Society* (1900; reprint, Chicago: University of Chicago Press, 1956), p. 3.

13. All of these networks and organizations have active websites.

14. National assessment expert Richard Stiggins, for instance, recently told Nebraska educators that the state is "leading the nation in assessment." Also, in 2002 the National Council on Measurement in Education honored the University of Nebraska's teacher education program for its innovative "assessment cohort" program, which awards graduate credit for studying and developing classroom assessments. That program is currently developing the country's first endorsement in assessment leadership. Finally, Nebraska's STARS system has received considerable attention in the pages of the *Kappan*: see Chris Gallagher, "A Seat at the Table," *Phi Delta Kappan*, March 2000, pp. 502-7; and Pat Roschewski, "Nebraskans Reach for the STARS," *Phi Delta Kappan*, April 2001, pp. 611-15.

15. See Roschewski, op. cit.

16. Nebraska assessment results are available at <http://reportcard.nde.state.ne.us>.

17. Chris Gallagher, "Charting STARS: The State of Assessment in the State of Nebraska," Year One Report, STARS Research and Evaluation Study, August 2002, executive summary available at [www.nde.state.ne.us/stars/resources/execsummary0101.pdf](http://www.nde.state.ne.us/stars/resources/execsummary0101.pdf); see also Chris Gallagher, "Charting STARS: Sustainability as Opportunity and Challenge," Year Two Report, STARS Comprehensive Evaluation Project, August 2003, available at [www.nde.state.ne.us/stars/index.html](http://www.nde.state.ne.us/stars/index.html).

18. This useful formulation comes from Darling-Hammond, p. 66.

19. On learning organizations as a key to sustainable reform, see Lauren B. Resnick and Megan Williams Hall, "Learning Organizations for Sustainable Education Reform," *Daedalus*, Fall 1998, pp. 89-118.

20. For a discussion of commitment and compliance, see Elaine Martin, *Changing Academic Work: Developing the Learning University* (Buckingham, U.K.: Society for Research into Higher Education and Open University Press, 1999).

21. Alfie Kohn, *The Case Against Standardized Testing* (Portsmouth, N.H.: Heinemann, 2000).

22. See, for instance, Michael Apple, *Educating the 'Right' Way* (New York: RoutledgeFalmer, 2001).

23. Pat Roschewski, "A Study of Effective Assessment Practices, Processes, and Procedures" (Doctoral dissertation, University of Nebraska, 2002).

24. Beverly Falk, *The Heart of the Matter* (Portsmouth, N.H.: Heinemann, 2000).

25. Paul Black and Dylan Wiliam, "Inside the Black Box: Raising Standards Through Classroom Assessment," *Phi Delta Kappan*, October 1998, pp. 139-48.

26. Richard J. Stiggins, "Assessment Crisis: The Absence of Assessment FOR Learning," *Phi Delta Kappan*, June 2002, pp. 758-65. **■**